

ЗАКЛЮЧЕНИЕ ДИССЕРТАЦИОННОГО СОВЕТА Д 003.011.01

НА БАЗЕ ФЕДЕРАЛЬНОГО ГОСУДАРСТВЕННОГО БЮДЖЕТНОГО
НАУЧНОГО УЧРЕЖДЕНИЯ «ФЕДЕРАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
ЦЕНТР ИНСТИТУТ ЦИТОЛОГИИ И ГЕНЕТИКИ СИБИРСКОГО
ОТДЕЛЕНИЯ РОССИЙСКОЙ АКАДЕМИИ НАУК»

ПО ДИССЕРТАЦИИ НА СОИСКАНИЕ УЧЕНОЙ СТЕПЕНИ
ДОКТОРА БИОЛОГИЧЕСКИХ НАУК

Аттестационное дело № _____

Дата защиты 25 октября 2017 г. протокол № 31

О присуждении Пономаренко Михаилу Павловичу
ученой степени доктора биологических наук.

Диссертация Пономаренко М. П. “Компьютерный анализ контекстно-зависимых количественных характеристик специфической биологической активности сайтов в составе геномной ДНК” по специальности 03.01.09 – математическая биология, биоинформатика, принята к защите 28.06.2017 г, протокол № 25, диссертационным советом Д 003.011.01 на базе Федерального государственного бюджетного научного учреждения «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук», (630090, Новосибирск, пр. акад. Лаврентьева, 10). Диссертационный совет Д 003.011.01 утвержден ВАК 15.01.2010, приказ ВАК № 1-7 и переутвержден Министерством образования и науки РФ 11.04.2012 года, приказ № 105/нк.

Соискатель: Пономаренко Михаил Павлович, 1962 года рождения, окончил Новосибирский государственный университет по специальности «Физика» в 1985 году, в 1994 г. защитил кандидатскую диссертацию на тему: «Исследование генетических текстов с помощью компьютерной технологии "Site-Video"» и получил степень кандидата биологических наук по

специальности генетика (диплом КН № 0006834 от 01 июля 1994 г.). Работает в ИЦиГ СО РАН с 1983 года, в настоящее время в должности старшего научного сотрудника лаборатории эволюционной биоинформатики и теоретической генетики Федерального государственного бюджетного научного учреждения «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук».

Диссертация выполнена в лаборатории эволюционной биоинформатики и теоретической генетики Федерального государственного бюджетного научного учреждения «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук».

Научный консультант: **Колчанов Николай Александрович**, доктор биологических наук, академик РАН, профессор, директор Федерального государственного бюджетного научного учреждения «Федеральный исследовательский центр Институт цитологии и генетики Сибирского отделения Российской академии наук», г. Новосибирск.

Официальные оппоненты:

1. **Макеев Всеволод Юрьевич**, член-корреспондент РАН, доктор физико-математических наук, заведующий отделом вычислительной системной биологии, ФГБУН «Институт общей генетики РАН им. Н. И. Вавилова», г. Москва
2. **Дубина Михаил Владимирович**, академик РАН, доктор медицинских наук, профессор, руководитель отдела молекулярно-генетических и нанобиологических технологий, Научно-исследовательский центр, ФГБОУ ВО ПСПбГМУ им. И. П. Павлова Минздрава России, г. Санкт-Петербург
3. **Бажан Сергей Иванович**, доктор биологических наук, заведующий теоретическим отделом, ФБУН ГНЦ вирусологии и биотехнологии «Вектор» Роспотребнадзора, р.п. Кольцово Новосибирской обл.

Оппоненты дали положительные отзывы на диссертацию.

Ведущая организация: Федеральное государственное учреждение «Федеральный исследовательский центр «Фундаментальные основы биотехнологии» Российской академии наук» (ФИЦ биотехнологии РАН), г. Москва в своем положительном заключении, подписанном профессором, д. б. н. Прохорчук Е. Б. и утвержденном директором Федерального государственного учреждения «Федеральный исследовательский центр «Фундаментальные основы биотехнологии» Российской академии наук», член-корреспондентом РАН, д. х. н. Поповым В. О., указало, что «Диссертация Пономаренко М. П. является научно-квалификационной работой, в которой на основании выполненных автором исследований разработаны теоретические положения о принципах ДНК-белковых взаимодействий, совокупность которых можно квалифицировать как научное достижение, что соответствует требованиям п. 9 «Положения о присуждении ученых степеней», утверждённого постановлением Правительства Российской Федерации от 24 сентября 2013 г. № 842, предъявляемым к диссертациям на соискание ученой степени доктора наук, а ее автор заслуживает присуждения искомой ученой степени по специальности 03.01.09 «Математическая биология, биоинформатика». Отзыв обсужден на межлабораторном семинаре ФИЦ биотехнологии РАН (протокол № 21 от 19 сентября 2017 года).»

Соискатель имеет всего 196 научных трудов из них 55 по теме диссертации, общим объемом 529 страниц, в том числе 30 статей, опубликованных в научных рецензируемых изданиях из Перечня ВАК (все индексированы в РИНЦ, Scopus и Web of Science), 11 тезисов в материалах всероссийских и международных конференций и 8 глав в монографиях.

Наиболее значительные статьи Пономаренко М. П. по теме диссертации:

1. **Ponomarenko, M.P.** Generating programs for predicting the activity of functional sites. / M.P. Ponomarenko, A.N. Kolchanova, N.A. Kolchanov // *J. Comput. Biol.* - 1997. - V. 4, N. 1. - P. 83 - 90.
2. **Пономаренко, М.П.** Предпочтительность ResA-филамента к последовательностям ДНК коррелирует с генетическим кодом. / М.П. Пономаренко, Ю.В. Пономаренко, И.И. Титов, Н.А. Колчанов, А.В. Мазин, С. Ковальчиковски // *ДАН.* – 1998. - Т. 363, № 1. - С. 122 - 125.
3. **Ponomarenko, M.P.** Identification of sequence-dependent features correlating to activity of DNA sites interacting with proteins. / M.P. Ponomarenko, J.V. Ponomarenko, A.S. Frolov, N.L. Podkolodny, L.K. Savinkova, N.A. Kolchanov, G.C. Overton // *Bioinformatics.* - 1999. - V. 15, N. 7/8. - P. 687 - 703
4. Vasiliev, G.V. Point mutations within 663-666 bp of intron 6 of the human TDO2 gene, associated with a number of psychiatric disorders, damage the YY1 transcription factor binding site. / G.V. Vasiliev, V.M. Merkulov, V.F. Kobzev, T.I. Merkulova, **M.P. Ponomarenko**, N.A. Kolchanov // *FEBS Lett.* – 1999. - V. 462, N. 1/2. - P. 85 - 88.
5. Ponomarenko, J.V. ACTIVITY: a database on DNA/RNA sites activity adapted to apply sequence-activity relationships from one system to another. / J.V. Ponomarenko, D.P. Furman, A.S. Frolov, N.L. Podkolodny, G.V. Orlova, **M.P. Ponomarenko**, N.A. Kolchanov, A. Sarai // *Nucleic Acids Res.* - 2001. - V. 29, N. 1. - P. 284 - 287.
6. Kirpota, O.O. Thermodynamic and kinetic basis for recognition and repair of 8-oxoguanine in DNA by human 8-oxoguanine-DNA glycosylase. / O.O. Kirpota, A.V. Endutkin, **M.P. Ponomarenko**, P.M. Ponomarenko, D.O. Zharkov, G.A. Nevinsky // *Nucleic Acids Res.* - 2011. - V. 39, N. 11. - P. 4836 - 4850.
7. Втюрина, Н.Н. Контекстные характеристики ДНК, значимые для ее повреждения ультрафиолетовым лазерным излучением с длиной волны 193 нм. / Н.Н. Втюрина, С.Л. Гроховский, А.Б. Васильев, И.И. Титов, П.М. Пономаренко, **М.П. Пономаренко**, С.Е. Пельтек, Ю.Д. Нечипуренко, Н.А. Колчанов // *ДАН.* – 2012. - Т. 447, № 2. - С. 217 - 222.

8. Savinkova, L.K. An experimental verification of the predicted effects of promoter TATA-box polymorphisms associated with human diseases on interactions between the TATA boxes and TATA-binding protein. / L.K. Savinkova, I.A. Drachkova, T.V. Arshinova, P.M. Ponomarenko, **M.P. Ponomarenko**, N.A. Kolchanov // PLoS ONE. - 2013. - V. 8, N. 2. - P. e54626.

На автореферат диссертации поступило 5 отзывов, все положительные. Отзывы прислали:

1. **Михаил Сергеевич Гельфанд**, доктор биологических наук, профессор, заместитель директора по научным вопросам ФГБУН «Институт проблем передачи информации им. А.А. Харкевича РАН», Москва. **Замечания:** “На стр. 20 автор выдвигает эволюционную гипотезу о возникновении гистон-подобных белков бактерий (в частности, HU) в ответ на ультрафиолетовое излучение. Представляется естественным проверить эту гипотезу сравнительно-геномными методами: сопоставить филетические паттерны этих белков с местами обитания бактерий. Современные базы данных (в частности, COG) позволят провести такой анализ силами одного хорошего студента. Более тонкое исследование заключалось бы в сопоставлении филогенетических деревьев данных белков с деревом бактерий, идентификации возможных горизонтальных переносов и определении времени и места возникновения этих белков (были ли они у LUCA?).

В ряде конкретных исследований большая корреляция между предсказанными и измеренными величинами определяется уникальными, отстоящими далеко от остальных, точками (рис. 12а, рис. 15а). Это стоило бы как-то прокомментировать.

Автор рассматривает очень большое количество параметров. Производились ли поправки на множественное тестирование при оценке статистической значимости?

Вообще, автор приводит только результаты успешных предсказаний. Это «ошибка выжившего» — известная проблема всех метаанализов, каковым, по сути, и является диссертация. Было бы весьма интересно узнать, были ли задачи, в которых предлагаемые авторами методы не сработали, сколько, в чем они заключались. Это позволило бы более точно оценить предлагаемые алгоритмы и конкретные формулы, показало границы применимости подхода и, с другой стороны, позволило бы наметить пути дальнейшего развития.

Сложностью для всех больших, продолжительных биоинформатических проектов является быстрое совершенствование экспериментальных методов: с таким трудом полученные результаты быстро устаревают. Так, очень естественным для данной работы было бы сопоставление результатов по предсказанию сайтов связывания транскрипционных факторов с данными ChIP-Seq, а исследование мутационных контекстов — с данными по секвенированию раковых опухолей (отметим, что эта область сейчас активно развивается).”

2. **Летягин Андрей Юрьевич**, доктор медицинских наук, руководитель Научно-исследовательского института клинической и экспериментальной лимфологии – филиала ФГБНУ «Федеральный исследовательский центр Институт цитологии и генетики СО РАН», г. Новосибирск.
3. **Дейнеко Игорь Владимирович**, кандидат биологических наук, группа по изучению биомаркеров инфекций и иммунитета, Центр Гельмгольца инфекционных исследований, г. Брауншвейг, Германия.
4. **Ратушняк Александр Савельевич**, доктор биологических наук, заведующий лабораторией биомедицинской информатики ФГБУН «Институт вычислительных технологий СО РАН», г. Новосибирск.
5. **Витяев Евгений Евгеньевич**, доктор физико-математических наук, ведущий научный сотрудник лаборатории теории вычислимости и прикладной логики, ФГБУН «Институт математики им. С.Л. Соболева СО РАН», г. Новосибирск. **Замечание:** “К сожалению, из автореферата

остается неясным, как на основе имеющихся экспериментальных данных соискатель формировал обучающие и контрольные выборки, соответственно, для компьютерного анализа и для оценки достоверности результатов этого анализа”.

Выбор официальных оппонентов и ведущей организации обосновывается тем, что официальные оппоненты - это высокоавторитетные специалисты по математической биологии и биоинформатики, публикующие свои результаты в ведущих международных и российских научных журналах по биологии и давшие свое письменное согласие выступить в качестве оппонентов. Ведущая организация является одним из лидеров среди Федеральных Исследовательских Центров в России в области математической биологии и биоинформатики.

Диссертационный совет отмечает, что на основании выполненных соискателем исследований предложен оригинальный подход к анализу последовательностей геномной ДНК в рамках теории аддитивной полезности для принятия решений и теории нечетких множеств, позволивший выявить качественно новые закономерности, отражающие взаимосвязь между уровнем функциональной активности сайтов и контекстно-зависимыми свойствами двойной спирали ДНК.

Разработаны алгоритмы и компьютерные системы для выявления контекстно-зависимых свойств спирали ДНК сайтов, достоверно коррелирующих с уровнем их биологической активности и адекватно отражающих такие биологически значимые характеристики, как сродство транскрипционных факторов к промоторам генов, частоты повреждения ДНК мутагенами и эффективность репарации этих повреждений.

Доказано, что использование контекстно-зависимых характеристик ДНК-сайтов повышает точность теоретических прогнозов, позволяя перейти от качественного распознавания наличия или отсутствия определенной

активности заданного сайта к количественной оценке этой активности, и расширяет границы применения биоинформатических предсказаний до решения задач ген-ориентированной селекции организмов, биотехнологии и персонализированной медицины.

Теоретическая значимость исследования обоснована тем, что применительно к проблематике диссертации эффективно, с получением имеющих новизну результатов, использованы теория полезности для принятия решений и теория нечетких множеств, позволившие единообразно проанализировать более миллиона вариантов конформационных и физико-химических свойств спирали ДНК функциональных сайтов и отобрать среди них те характеристики, которые в наибольшей мере удовлетворяют критериям применимости регрессионного анализа к биоинформатическим прогнозам на выборках имеющих экспериментальных данных.

С помощью алгоритмов, разработанных в диссертации, **проведена** оценка величин сродства ТАТА-связывающего белка, Cro-репрессора, активатора CRP и транскрипционных факторов USF, MEF2, YY1 к сайтам их связывания в геномах, а также частот повреждения гуанинов в геномной ДНК при воздействии ультрафиолетового излучения лазера с длиной волны 193 нм, эффективности связывания белка RecA с одонитевой ДНК, константы Михаэлиса и каталитической константы 8-оксогуанин-ДНК гликозилазы OGG1 человека. Теоретические предсказания подтверждены с использованием независимых экспериментальных данных.

Доказано, что замены 663G→A и 666G→T в интроне 6 гена *TDO2* человека повреждают сайт связывания транскрипционного фактора YY1, что клинически проявляется в форме поведенческих расстройств.

Доказано, что транскрипционные факторы кластеризуются в две группы по сходству спиралей ДНК сайтов их связывания. Одна группа включает белки с β-слоем и с гомеодоменом, которые взаимодействуют с ДНК по принципу “ключ-замок”; другая - основные и Zn-координируемые белки, взаимодействие которыми с ДНК обеспечивают локальные избытки положительного и

отрицательного электростатических зарядов, соответственно.

Значение полученных соискателем результатов исследования для практики подтверждается тем, что разработана компьютерная система bDNAvideo, включающая базу данных по контекстно-зависимым конформационным и физико-химическим свойствам спирали ДНК, о практическом использовании которой в разработке компьютерных систем SITECON (Россия), DISCOVER (США), BiDaS (Греция), CRoSSeD (Бельгия), FeatureScan, DiProDB, BioBayesNet и ProMapper (все из Германии) было сообщено в соответствующих публикациях авторами этих систем.

Разработана компьютерная система Activity для анализа последовательностей сайтов ДНК с экспериментально измеренными величинами активности на основе использования оригинальной линейно-аддитивной регрессионной математической модели, оценивающей величины заданной активности функциональных сайтов ДНК по их нуклеотидному контексту, с учетом встречаемости нуклеотидов, консенсуса и позиционно-весовой матрицы, а также контекстно-зависимых физико-химических и конформационных свойств двойной спирали ДНК. Эффективность этой системы подтверждена результатами независимых экспериментов.

Полученные соискателем теоретические знания используются в научно-исследовательских учреждениях, изучающих закономерности биологических процессов методами математической биологии и биоинформатики (Россия, США, Греция, Бельгия, Германия), а также в образовательном процессе при подготовке специалистов в рамках учебного плана курсов «Введение в информационную биологию» и «Современные проблемы биоинформатики и системной биологии» кафедры информационной биологии факультета естественных наук Новосибирского государственного университета.

Оценка достоверности результатов исследования выявила, что регрессионные уравнения для оценки величин специфической биологической активности сайтов ДНК построены на основе литературных и оригинальных опубликованных экспериментальных данных по измерению этой активности,

а количественные величины биологической активности сайтов, вычисленные с использованием этих уравнений, статистически достоверно коррелируют с соответствующими величинами, измеренными в независимых экспериментах.

Использованы современные методики сбора и обработки исходных данных, представленных репрезентативными выборочными совокупностями нуклеотидных последовательностей с экспериментально измеренными для них величинами специфической биологической активностей сайтов в составе геномной ДНК.

Установлено достоверное соответствие между результатами компьютерного анализа последовательностей ДНК сайтов, представленных в диссертационной работе, и экспериментальными данными из независимых источников.

Личный вклад автора состоит в разработке компьютерных систем bDNAvideo и Activity для выявления физико-химических и конформационных характеристик последовательностей сайтов ДНК, определяющих их биологическую активность; компьютерном моделировании и анализе влияния контекста сайта на его функциональную активность; построении регрессионных уравнений с использованием выборок нуклеотидных последовательностей сайтов с экспериментально измеренными для них величинами специфической активности; а также участии в планировании контрольных экспериментов по верификации теоретически предсказанных результатов, их интерпретации, апробации на научных конференциях и подготовке публикаций. Экспериментальные исследования, представленные в статьях по теме диссертации, выполнены соавторами.

Полученные соискателем научные результаты соответствуют п. 3. “Компьютерная геномика, протеомика, иммуномика”, п. 8 “Математические модели, численные методы и программные средства применительно к процессам получения, накопления, обработки и систематизации биологических и медицинских данных и знаний”, п. 9 “Организация, ведение и использование автоматизированных банков данных по биологии и медицине,

в т.ч. банков междисциплинарных данных” и п. 10 “Интеллектуальные системы анализа и прогнозирования свойств биологических объектов на основе специализированных баз и банков данных и знаний (в т.ч. полнотекстовых)” паспорта специальностей 03.01.09 – математическая биология, биоинформатика (биологические науки).

Диссертационным советом сделан вывод о том, что диссертация М.П. Пономаренко представляет собой законченную научно-квалификационную работу, которая соответствует критериям пункта 9 «Положения о присуждении ученых степеней», утвержденным постановлением Правительства Российской Федерации от 24 сентября 2013 г № 842.

На заседании 25 октября 2017 года диссертационный совет принял решение присудить Пономаренко Михаилу Павловичу учёную степень доктора биологических наук.

При проведении тайного голосования диссертационный совет в количестве 19 человек, из них 7 докторов наук по специальности, участвующих в заседании, из 27 человек, входящих в состав совета, проголосовали: за – 19, против – нет, недействительных бюллетеней – нет.

Председатель

диссертационного совета

академик РАН



Ученый секретарь

диссертационного совета,

доктор биологических наук

В.К. Шумный

Т.М. Хлебодарова

25.10.2017 г.